Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 25 13:48:45 2000; MasPar time 24.38 Seconds
834.844 Million cell updates/sec

Tabular output not generated.

Title: >US-08-956-991-11
Description: (1-1571) from US08956991A.pep
Perfect Score: 11189
Sequence: 1 MWILALSLFQSFANVFSEDL.....HSAGCAEKQAKEAARCKEFS 1571

Scoring table: PAM 150
Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfiles1

Statistics: Mean 38.265; Variance 226.750; scale 0.169

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2081	18.6	465	2	US-08-752-	Sequence 5, Applicatio	1.19e-136
2	2071	18.5	462	2	US-08-752-	Sequence 7, Applicatio	6.27e-136
3	740	6.6	1447	3	PCT-US94-0	Sequence 2, Applicatio	8.27e-41
4	629	5.6	1018	1	US-08-408-	Sequence 6, Applicatio	4.68e-33
5	629	5.6	1018	1	US-08-408-	Sequence 6, Applicatio	4.68e-33
6	629	5.6	1018	1	US-08-714-	Sequence 6, Applicatio	4.68e-33
7	606	5.4	1018	1	US-08-452-	Sequence 2, Applicatio	1.86e-31
8	569	5.1	1911	1	US-08-348-	Sequence 5, Applicatio	6.80e-29
9	569	5.1	1911	2	US-08-800-	Sequence 5, Applicatio	6.80e-29
10	569	5.1	1911	3	PCT-US94-1	Sequence 5, Applicatio	6.80e-29
11	486	4.3	1501	2	US-08-716-	Sequence 3, Applicatio	3.52e-23
12	486	4.3	1501	2	US-08-447-	Sequence 3, Applicatio	3.52e-23
13	449	4.0	596	2	US-08-752-	Sequence 13, Applicati	1.18e-20
14	448	4.0	612	2	US-08-752-	Sequence 11, Applicati	1.38e-20
15	391	3.5	607	2	US-08-752-	Sequence 12, Applicati	1.00e-16
16	381	3.4	605	2	US-08-752-	Sequence 8, Applicatio	4.71e-16
17	379	3.4	946	3	PCT-US95-0	Sequence 13, Applicati	6.42e-16
18	361	3.2	615	2	US-08-752-	Sequence 9, Applicatio	1.03e-14
19	359	3.2	630	2	US-08-752-	Sequence 14, Applicati	1.40e-14
20	322	2.9	478	3	PCT-US95-0	Sequence 15, Applicati	4.03e-12
21	322	2.9	860	3	PCT-US95-0	Sequence 19, Applicati	4.03e-12
22	322	2.9	868	3	PCT-US95-0	Sequence 21, Applicati	4.03e-12
23	320	2.9	868	1	US-08-374-	Sequence 1, Applicatio	5.46e-12
24	320	2.9	868	2	US-08-644-	Sequence 1, Applicatio	5.46e-12
25	326	2.9	869	1	US-08-374-	Sequence 16, Applicati	2.19e-12
26	326	2.9	869	2	US-08-644-	Sequence 29, Applicati	2.19e-12
27	282	2.5	2231	1	US-08-153-	Sequence 16, Applicati	1.70e-09
28	262	2.3	2324	1	US-08-283-	Sequence 1, Applicatio	3.35e-08
29	262	2.3	2324	3	PCT-US95-0	Sequence 1, Applicatio	3.35e-08
30	262	2.3	2327	4	5455158-1	Patent No. 5455158.	3.35e-08
31	262	2.3	2386	2	US-09-016-	Sequence 12, Applicati	3.35e-08
32	262	2.3	2446	2	US-08-551-	Sequence 2, Applicatio	3.35e-08
33	262	2.3	2446	3	PCT-US93-1	Sequence 2, Applicatio	3.35e-08
34	233	2.1	252	2	US-08-414-	Sequence 57, Applicati	2.39e-06
35	235	2.1	287	2	US-08-414-	Sequence 49, Applicati	1.79e-06
36	235	2.1	310	2	US-08-414-	Sequence 45, Applicati	1.79e-06
37	235	2.1	315	2	US-08-414-	Sequence 47, Applicati	1.79e-06
38	235	2.1	338	2	US-08-414-	Sequence 43, Applicati	1.79e-06
39	235	2.1	338	2	US-08-414-	Sequence 42, Applicati	1.79e-06
40	234	2.1	419	1	5169835-2	Patent No. 5169835.	2.07e-06
41	232	2.1	549	2	US-08-836-	Sequence 11, Applicati	2.77e-06
42	232	2.1	574	2	US-08-836-	Sequence 21, Applicati	2.77e-06
43	234	2.1	828	1	US-08-261-	Sequence 2, Applicatio	2.07e-06
44	229	2.0	287	2	US-08-414-	Sequence 48, Applicati	4.29e-06
45	229	2.0	304	2	US-08-414-	Sequence 44, Applicati	4.29e-06

ALIGNMENTS

```
RESULT    1
ID   US-08-752-307B-5     STANDARD;     PRT;   465 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 5, Application US/08752307B
XX
CC    Sequence 5, Application US/08752307B
CC    Patent No. 5952171
CC     GENERAL INFORMATION:
CC       APPLICANT:  McCarthy, Sean A.
CC       APPLICANT:  Gearing, David P.
CC       APPLICANT:  Levinson, Douglas A.
CC       TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENES
CC       TITLE OF INVENTION:  ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Fish & Richardson, P.C.
CC         STREET:  225 Franklin Street
CC         CITY:  Boston
CC         STATE:  MA
CC         COUNTRY:  US
CC         ZIP:  02110-2804
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  Windows95
CC         SOFTWARE:  FastSEQ for Windows Version 2.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/752,307B
CC         FILING DATE:  19-NOV-1996
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Meiklejohn, Ph.D., Anita L.
CC         REGISTRATION NUMBER:  35,283
CC         REFERENCE/DOCKET NUMBER:  09404/020001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  617-542-5070
CC         TELEFAX:  617-542-8906
CC         TELEX:  200154
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
```

```
CC         LENGTH:  465 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FRAGMENT TYPE:  internal
SQ   SEQUENCE   465 AA; 51210 MW; 1158127 CN;

  Query Match           18.6%;  Score 2081;  DB 2;  Length 465;
  Best Local Similarity  57.1%;  Pred. No. 1.19e-136;
  Matches   265;  Conservative  105;  Mismatches  93;  Indels   1;  Gaps   1;

Db       1 MWLVTFLLLLDSLHKARPEDVGTSLYFVNDSLQQVTFSSSVGVVVPCPAAGSPSAALRWY 60
           ||::: | |::|: :  :||: :|||||| |||:| |:|: | :|||||| |: :|||| 
Qy       1 MWILA-LSLFQSFANVFSEDLHSSLYFVNASLQEVVFASTTGTLVPCPAAGIPPVTLRWY 59

Db      61 LATGDDIYDVPHIRHVHANGTLQLYPFSPSAFNSFIHDNDYFCTAENAAGKIRSPNIRVK 120
           ||||::|||| |||||:|||||::||:||:|:::|||| |:|||||::||||| :::|
Qy      60 LATGEEIYDVPGIRHVHPNGTLQIFPFPPSSFSTLIHDNTYYCTAENPSGKIRSQDVHIK 119

Db     121 AVFREPYTVRVEDQRSMRGNVAVFKCLIPSSVQEYVSVVSWEKDTVSIIPENRFFITYHG 180
           ||:||||||||||::|||||||||:|||||: |::|||||||||||:: :|||:||  |
Qy     120 AVLREPYTVRVEDQKTMRGNVAVFKCIIPSSVEAYITVVSWEKDTVSLVSGSRFLITSTG 179

Db     181 GLYISDVQKEDALSTYRCITKHKYSGETRQSNGARLSVTDPAESIPTILDGFHSQEVWAG 240
           :||| |||:||:|  ||||||:|:|:|||||||:||| |:|||:| |:|||||  :   ||
Qy     180 ALYIKDVQNEDGLYNYRCITRHRYTGETRQSNSARLFVSDPANSAPSILDGFDHRKAMAG 239

Db     241 HTVELPCTASGYPIPAIRWLKDGRPLPADSRWTKRITGLTISDLRTEDSGTYICEVTNTF 300
           : ||||| | | | |  |||||  ||   :|  | :||| | ::|  |||:|:|||:| :
Qy     240 QRVELPCKALGHPEPDYRWLKDNMPLELSGRFQKTVTGLLIENIRPSDSGSYVCEVSNRY 299

Db     301 GSAEATGILMVIDPLHVTLTPKKLKTGIGSTVILSCALTGSPEFTIRWYRNTELVLPDEA 360
           |:|   | | | :||  |::|:|:|:::|| | |||::||: :  : |||| |:: |
Qy     300 GTAKVIGRLYVKQPLKATISPRKVKSSVGSQVSLSCSVTGTEDQELSWYRNGEILNPGKN 359

Db     361 ISIRGLSNETLLITSAQKSHSGAYQCFATRKAQTAQDFAIIALEDGTPRIVSSFSEKVVN 420
           : | |:::| |::    || :|||||  :   :|||:  : ||||||:|:|:||||||:
Qy     360 VRITGINHENLIMDHMVKSDGGAYQCFVRKDKLSAQDYVQVVLEDGTPKIISAFSEKVVS 419

Db     421 PGEQFSLMCAAKGAPPPTVTWALDDEPIVRDGSHRTNQYTMSDG 464
           |:|  ||||  ||:| ||:||:|||:||:: |||| :|   |:|
Qy     420 PAEPVSLMCNVKGTPLPTITWTLDDDPILKGGSHRISQMITSEG 463


RESULT    2
ID   US-08-752-307B-7     STANDARD;     PRT;   462 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 7, Application US/08752307B
XX
CC    Sequence 7, Application US/08752307B
CC    Patent No. 5952171
CC     GENERAL INFORMATION:
CC       APPLICANT:  McCarthy, Sean A.
CC       APPLICANT:  Gearing, David P.
CC       APPLICANT:  Levinson, Douglas A.
CC       TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENES
CC       TITLE OF INVENTION:  ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Fish & Richardson, P.C.
CC         STREET:  225 Franklin Street
CC         CITY:  Boston
CC         STATE:  MA
CC         COUNTRY:  US
CC         ZIP:  02110-2804
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  Windows95
CC         SOFTWARE:  FastSEQ for Windows Version 2.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/752,307B
CC         FILING DATE:  19-NOV-1996
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Meiklejohn, Ph.D., Anita L.
CC         REGISTRATION NUMBER:  35,283
CC         REFERENCE/DOCKET NUMBER:  09404/020001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  617-542-5070
CC         TELEFAX:  617-542-8906
CC         TELEX:  200154
CC     INFORMATION FOR SEQ ID NO:  7:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  462 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   462 AA; 50937 MW; 1144658 CN;

  Query Match           18.5%;  Score 2071;  DB 2;  Length 462;
  Best Local Similarity  57.4%;  Pred. No. 6.27e-136;
  Matches   263;  Conservative  104;  Mismatches  90;  Indels   1;  Gaps   1;

Db       1 MWLVTFLLLLDSLHKARPEDVGTSLYFVNDSLQQVTFSSSVGVVVPCPAAGSPSAALRWY 60
           ||::: | |::|: :  :||: :|||||| |||:| |:|: | :|||||| |: :|||| 
Qy       1 MWILA-LSLFQSFANVFSEDLHSSLYFVNASLQEVVFASTTGTLVPCPAAGIPPVTLRWY 59

Db      61 LATGDDIYDVPHIRHVHANGTLQLYPFSPSAFNSFIHDNDYFCTAENAAGKIRSPNIRVK 120
           ||||::|||| |||||:|||||::||:||:|:::|||| |:|||||::||||| :::|
Qy      60 LATGEEIYDVPGIRHVHPNGTLQIFPFPPSSFSTLIHDNTYYCTAENPSGKIRSQDVHIK 119

Db     121 AVFREPYTVRVEDQRSMRGNVAVFKCLIPSSVQEYVSVVSWEKDTVSIIPENRFFITYHG 180
           ||:||||||||||::|||||||||:|||||: |::|||||||||||:: :|||:||  |
Qy     120 AVLREPYTVRVEDQKTMRGNVAVFKCIIPSSVEAYITVVSWEKDTVSLVSGSRFLITSTG 179

Db     181 GLYISDVQKEDALSTYRCITKHKYSGETRQSNGARLSVTDPAESIPTILDGFHSQEVWAG 240
           :||| |||:||:|  ||||||:|:|:|||||||:||| |:|||:| |:|||||  :   ||
Qy     180 ALYIKDVQNEDGLYNYRCITRHRYTGETRQSNSARLFVSDPANSAPSILDGFDHRKAMAG 239

Db     241 HTVELPCTASGYPIPAIRWLKDGRPLPADSRWTKRITGLTISDLRTEDSGTYICEVTNTF 300
           : ||||| | | | |  |||||  ||   :|  | :||| | ::|  |||:|:|||:| :
Qy     240 QRVELPCKALGHPEPDYRWLKDNMPLELSGRFQKTVTGLLIENIRPSDSGSYVCEVSNRY 299

Db     301 GSAEATGILMVIDPLHVTLTPKKLKTGIGSTVILSCALTGSPEFTIRWYRNTELVLPDEA 360
           |:|   | | | :||  |::|:|:|:::|| | |||::||: :  : |||| |:: |
Qy     300 GTAKVIGRLYVKQPLKATISPRKVKSSVGSQVSLSCSVTGTEDQELSWYRNGEILNPGKN 359

Db     361 ISIRGLSNETLLITSAQKSHSGAYQCFATRKAQTAQDFAIIALEDGTPRIVSSFSEKVVN 420
           : | |:::| |::    || :|||||  :   :|||:  : ||||||:|:|:||||||:
Qy     360 VRITGINHENLIMDHMVKSDGGAYQCFVRKDKLSAQDYVQVVLEDGTPKIISAFSEKVVS 419

Db     421 PGEQFSLMCAAKGAPPPTVTWALDDEPIVRDGSHRTNQ 458
           |:|  ||||  ||:| ||:||:|||:||:: |||| :|
Qy     420 PAEPVSLMCNVKGTPLPTITWTLDDDPILKGGSHRISQ 457


RESULT    3
ID   PCT-US94-05277-2     STANDARD;     PRT;   1447 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application PC/TUS9405277
XX
CC    Sequence 2, Application PC/TUS9405277
CC     GENERAL INFORMATION:
CC       APPLICANT:  Bruskin, Arthur
```

```
CC       APPLICANT:  Jarosz, David E.
CC       APPLICANT:  Johnson, Karen
CC       APPLICANT:  Kinzler, Kenneth W.
CC       APPLICANT:  Vogelstein, Bert
CC       APPLICANT:  Zabrecky, James R.
CC       TITLE OF INVENTION:  Antibodies Specific for DCC Gene Product
CC       NUMBER OF SEQUENCES:  2
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Banner, Birch, McKie & Beckett
CC         STREET:  1001 G Street, N.W.
CC         CITY:  Washington
CC         STATE:  D.C.
CC         COUNTRY:  USA
CC         ZIP:  20001
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/05277
CC         FILING DATE:
CC         CLASSIFICATION:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Kagan, Sarah A.
CC         REGISTRATION NUMBER:  32,141
CC         REFERENCE/DOCKET NUMBER:  01107.42709
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  202.508.9100
CC         TELEFAX:  202.508.9299
CC         TELEX:  197430 BBMB UT
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1447 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   1447 AA; 158455 MW; 11668854 CN;

  Query Match             6.6%;  Score 740;  DB 3;  Length 1447;
  Best Local Similarity  29.4%;  Pred. No. 8.27e-41;
  Matches   226;  Conservative  179;  Mismatches 295;  Indels  70;  Gaps  53;

Db     72 IKWKKDGIHLALGMDERKQQLSNGSLLIQNILHSRHHKPDEGLYQCEASLGDSGSIISRT 131
Qy    535 IKWYKNS-NL-LPFNHRQVAFENNGTL--K-LSDVQKEVDEGEYTCNV-LVQPQLSTSQS 588

Db    132 AKVAVAGPLRFLSQTESVTAF-MGDTVLLKCEVIGEPMP-TIHWQKNQQDLTPIPGDSRV 189
Qy    589 VHVTVKVP-PFIQPFE-FPRFSIGQRVFIPCVVVSGDLPITITWQKDGRPIPGSLGVTID 646

Db    190 VVLPSGALQISRLQPGDIGIYRCSARNPASSRTGNEAEVRILSDPGLHRQLYFLQRPSNV 249
Qy    647 NIDFTSSLRISNLSLMHNGNYTCIARNEAAA-VEHQSQL-IVRVP--PK--FVVQ-PRDQ 699

Db    250 VAIEGKDAVLECCVSGYPPPSFTW-L-RGEEVIQLR--S-K-KYSLLGGSNLLISNVTDD 303
Qy    700 DGIYGKAVILNCSAEGYPVPTIVWKFSKGAGVPQFQPIALNGRIQVLSNGSLLIKHVVEE 759

Db    304 DSGMYTCVVTYK-NENISASAELTVLVPPWFLNHPSNLYAYESMDIFECTVSG-KPVPT 361
Qy    760 DSGYYLCKVSNDVGADVSKSMYLTVKIPAMITSYPNTTLATQGQKKEMSCTAHGEKPI-I 818

Db    362 VNWMKNGDVVIPS-D-YF-QI--VGG---SNLRILGVVKSDEGFYQCVAENEAGNAQTSA 413
Qy    819 VRWEKEDRIINPEMARYLVSTKEVGEEVISTLQILPTVREDSGFFSCHAINSYGEDRGII 878

Db    414 QLIVPKPAIPSSSVLPSAPRDVVPVLVSSRFVRLSWRPPAEAKGNIQTFTVFFSREGDNR 473
Qy    879 QLTVQEP--PDP---PEI--EIKDV--KARTITLRWTMGFDGNSPITGYDIECKNKSDSW 929

Db    474 ERALNTTQPGS-LQ-LTVGNLKPEAMYTFRVVAYNEWGPGESSQPIKVATQPELQVPGPV 531
Qy    930 DSAQRTKDVSPQLNSATIIDIHPSSTYSIRMYAKNRIGKSEPSNELTITAD-EAAPDGPP 988

Db    532 ENLQAVSTSPTSILITWEPPA-Y-ANGPVQGYRLFCTEVSTGKE-Q-NI-EVD--GLS-- 582
Qy    989 QEVHLEPISSQSIRVTWKAPKKHLQNGIIRGYQIGYREYSTGGNFQFNIISVDTSGDSEV 1048

Db    583 YKLEGLKKFTEYSLRFLAYNRYGPGVSTDDITVVTLSDVPSAPPQNVSLEVVNSRSIKVS 642
Qy   1049 YTLDNLNKFTQYGLVVQACNRAGTGPSSQEIITTTLEDVPSYPPENVQAIATSPESISIS 1108

Db    643 WLPPPSGTQNGFITGYKIRHRKTTRRGEM-ETLE-PNNLWYL-FTGLEKGSQYSFQVSAM 699
Qy   1109 WSTLSKEALNGILQGFRVIYWANLMDGELGEIKNITTTQPSLELDGLEKYTNYSIQVLAF 1168

Db    700 TVNGIGPPSNWYTAETPENDLDESQVPDQPSSLHVRPQTNCII-MSWTPPLNPNIVVRGY 758
Qy   1169 TRAGDGVRSE--QIFTRTKE-D---VPGPPAGVKAAAASASMVFVSWLPPLKLNGIIRKY 1222

Db    759 IIGYGVGSPYAETVRV-DSK-QRY-YSIERLESSSHYVISLKAFNNAGEG 805
Qy   1223 TV-F-CSHPYPTVISEFEASPDSFSYRIPNLSRNRQYSVWVVAVTSAGRG 1270


RESULT    4
ID   US-08-408-093-6     STANDARD;     PRT;   1018 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 6, Application US/08408093
XX
CC    Sequence 6, Application US/08408093
CC    Patent No. 5688916
CC     GENERAL INFORMATION:
CC       APPLICANT:  Reid, Robert A.
CC       APPLICANT:  Hemperly, John J.
CC       TITLE OF INVENTION:  Human Cell Adhesion Molecule and Nucleic
CC       TITLE OF INVENTION:  Acid Sequences
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Richard J. Rodrick, Becton Dickinson
CC         ADDRESSEE:  and Company
CC         STREET:  One Becton Drive
CC         CITY:  Franklin Lakes
CC         STATE:  NJ
CC         COUNTRY:  USA
CC         ZIP:  07417
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/408,093
CC         FILING DATE:  21-MAR-1995
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/040,741
CC         FILING DATE:  26 MAR 1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Fugit, Donna R.
CC         REGISTRATION NUMBER:  32,135
CC         REFERENCE/DOCKET NUMBER:  P-2630
CC     INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1018 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Homo sapiens
CC       FEATURE:
```

```
CC       NAME/KEY:  Disulfide-bond
CC       LOCATION:  45..94
CC     FEATURE:
CC       NAME/KEY:  Disulfide-bond
CC       LOCATION:  138..191
CC     FEATURE:
CC       NAME/KEY:  Disulfide-bond
CC       LOCATION:  243..290
CC     FEATURE:
CC       NAME/KEY:  Disulfide-bond
CC       LOCATION:  332..371
CC     FEATURE:
CC       NAME/KEY:  Disulfide-bond
CC       LOCATION:  416..464
CC     FEATURE:
CC       NAME/KEY:  Disulfide-bond
CC       LOCATION:  506..563
CC     FEATURE:
CC       NAME/KEY:  Domain
CC       LOCATION:  604..657
CC       OTHER INFORMATION:  /label= FLR
CC       OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC       OTHER INFORMATION:  III-like repeat"
CC     FEATURE:
CC       NAME/KEY:  Domain
CC       LOCATION:  707..760
CC       OTHER INFORMATION:  /label= FLR
CC       OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC       OTHER INFORMATION:  III-like repeat"
CC     FEATURE:
CC       NAME/KEY:  Domain
CC       LOCATION:  809..857
CC       OTHER INFORMATION:  /label= FLR
CC       OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC       OTHER INFORMATION:  III-like repeat"
CC     FEATURE:
CC       NAME/KEY:  Domain
CC       LOCATION:  905..952
CC       OTHER INFORMATION:  /label= FLR
CC       OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC       OTHER INFORMATION:  III-like repeat"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  188
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  238
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  318
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  437
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  453
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  474
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  501
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  571
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
CC     FEATURE:
CC       NAME/KEY:  Modified-site
CC       LOCATION:  913
CC       OTHER INFORMATION:  /label= ASN-glycos
CC       OTHER INFORMATION:  /note= "potential site of ASN-linked
CC       OTHER INFORMATION:  glycosylation"
SQ   SEQUENCE   1018 AA; 113333 MW; 5493875 CN;

  Query Match            5.6%;  Score 629;  DB 1;  Length 1018;
  Best Local Similarity  25.0%;  Pred. No. 4.68e-33;
  Matches   198;  Conservative  196;  Mismatches 348;  Indels  49;  Gaps  33;

Db    241 PADIVVQFKDVYALMGQNVTLECFALGNPVPDIRWRKVLEPMP-STAEIS-TSGAVLKIF 298
          || |    |:: |: |::  : |  :| |   |:| |  : :| :  :::  : : ||:
Qy    504 PASIR-PMKNITAIAGRDTYIHCRVIGYPYYSIKWYKNSNLLPFNHRQVAFENNGTLKLS 562

Db    299 NIQLE-DEGIYECEA-ENIRGKDKHQARIYVQAFPEWVEHINDTEVDIGSDLYWPCVATG 356
          ::| | ||| | |:     :    :  :: |   |  :: ::     ||  :: ||| :
Qy    563 DVQKEVDEGEYTCNVLVQPQLSTSQSVHVTVKV-PPFIQPFEFPRFSIGQRVFIPCVVVS 621

Db    357 KPIP-TIRWLKNGYAYH-K-G------E----LRLYDVTFENAGMYQCIAENTYGAIYAN 403
            :| || | |:| :     |      :    ||: :::: : | | ||| |  :|:
Qy    622 GDLPITITWQKDGRPIPGSLGVTIDNIDFTSSLRISNLSLMHNGNYTCIARNEAAAVEHQ 681

Db    404 AELKILALAPTFEMNPMKKKILAAKGGRVIIECKPKAAPKPKFSW--SKGT---EWL-VN 457
          ::| |: ::| | : |     : :|:  ||::| : : | | : |  |||:   :   :
Qy    682 SQL-IVRVPPKFVVQPRDQDGIYGKA--VILNCSAEGYPVPTIVWKFSKGAGVPQFQPIA 738

Db    458 -SSRILIWEDGSLEINNITRNDGGIYTCFAENNRGKANSTGT-LVITDPTRIILAPINAD 515
           ::|| :  :||| |:::  :|:| | |   |: |   | :  | :  |: |   |  :
Qy    739 LNGRIQVLSNGSLLIKHVVEEDSGYYLCKVSNDVGADVSKSMYLTVKIPAMITSYPNTTL 798

Db    516 ITVGENATMQCAASFDPALDLTFVWSFNGYVIDFNKEN-IHYQRNFMLDSNGELLIRNAQ 574
          | |::  | |:|  : :: :   |  :  :|: ::   :   ::   :  : | |  :
Qy    799 ATQGQKKEMSCTAHGEKPI-IV-RWEKEDRIINPEMARYLVSTKEVGEEVISTLQILPTV 856

Db    575 LKHAGRYTCTAQTIVDNSSASADLVVRGPPGPPGGLRIEDIRATSVALTWSRGSDNHSPI 634
            :| ::| |     :  :  :| |: || ||  : | |::| :::| |: | | :|||
Qy    857 REDSGFFSCHAINSYGEDRGIIQLTVQEPPDPPE-IEIKDVKARTITLRWTMGFDGNSPI 915

Db    635 SKYTIQTKTILSDDWKDAKTDPPIIEGNMEAARAVDLIPWMEYEFRVVATNTLGRGEPSI 694
          : | |: |   || | |      :   :::|  :|: |   | :|: | | :|::|||
Qy    916 TGYDIECKNK-SDSWDSAQRTKDV-SPQLNSATIIDIHPSSTYSIRMYAKNRIGKSEPSN 973

Db    695 PSNRIKTDGAAPNVAPSDVGGGGGRNRELTITWAPLSREYHYGNNFGYIVAFKPFD-GEE 753
          | :| |||: :| :|     :: : :|| :  :  : | |  || :::: :  | :
Qy    974 ELT-ITADEAAPDGPPQEVHLEPISSQSIRVTWKAPKKHLQNGIIRGYQIGYREYSTGGN 1032

Db    754 WK-KVTVTNPDTGRYVHKDETMSPSTAFQVKVKAFNNKGDGPYSLLAVINSAQDAPSEAP 812
            ::   :      |   : ::  | : : | | |  | || |   :  : :| || :|
Qy   1033 FQFNIISVDTSGDSEVYTLDNLNKFTQYGLVVQACNRAGTGPSSQEIITTTLEDVPSYPP 1092

Db    813 TEVGVKVLSSSEISVHWE----HVLEKIVESYQIRYWAAHDKEEAANRVQVTSQEYSARL 868
           :|     |:  ||: |       |: |:::::: |||     | ::   :|: : |  |
Qy   1093 ENVQAIATSPESISISWSTLSKEALNGILQGFRVIYWANLMDGELGEIKNITTTQPSLEL 1152

Db    869 ENLLPDTQYFIEVGACNSAGCGPPSDMIEAFTKKAPPSQPPRIISSVRSGSRYIITWDHV 928
```

```
          : |   | | |:| |   || |  |: | : ||   |: |: : ::  |:|  :::|  :
Qy   1153 DGLEKYTNYSIQVLAFTRAGDGVRSEQIFTRTKEDVPGPPAGVKAAAASASMVFVSW--L 1210

Db    929 VALSNESTVTGYKVLY-RPDGQHDGKL-YSTHKHSIEVP-IPRDGEYVVEVRAHSDGGDG 985
           :|  :: :  | |:  :|     : :  |    |  :| ::|: :| | | | : :| |
Qy   1211 PPLKLNGIIRKYTVFCSHPYPTVISEFEASPDSFSYRIPNLSRNRQYSVWVVAVTSAGRG 1270

Db    986 VVSQVKISGAP 996
            |:: |:  |
Qy   1271 NSSEI-ITVEP 1280


RESULT    5
ID   US-08-408-420A-6     STANDARD;     PRT;   1018 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 6, Application US/08408420A
XX
CC    Sequence 6, Application US/08408420A
CC    Patent No. 5731154
CC     GENERAL INFORMATION:
CC       APPLICANT:  Reid, Robert A.
CC       APPLICANT:  Hemperly, John J.
CC       TITLE OF INVENTION:  Human Cell Adhesion Molecule and Nucleic
CC       TITLE OF INVENTION:  Acid Sequences
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Richard J. Rodrick, Becton Dickinson and
CC         ADDRESSEE:  Company
CC         STREET:  One Becton Drive
CC         CITY:  Franklin Lakes
CC         STATE:  NJ
CC         COUNTRY:  US
CC         ZIP:  07417
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/408,420A
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Fugit, Donna R.
CC         REGISTRATION NUMBER:  32,135
CC         REFERENCE/DOCKET NUMBER:  P-2630
CC      INFORMATION FOR SEQ ID NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1018 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Homo sapiens
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  45..94
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  138..191
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  243..290
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  332..371
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  416..464
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  506..563
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  604..657
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  707..760
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  809..857
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  905..952
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  188
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  238
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  318
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  437
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  453
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  474
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  501
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  571
CC         OTHER INFORMATION:  /label= ASN-glycos
```

```
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  913
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
SQ   SEQUENCE   1018 AA; 113333 MW; 5493875 CN;

  Query Match             5.6%;  Score 629;  DB 1;  Length 1018;
  Best Local Similarity  25.0%;  Pred. No. 4.68e-33;
  Matches   198;  Conservative  196;  Mismatches 348;  Indels  49;  Gaps  33;

Db    241 PADIVVQFKDVYALMGQNVTLECFALGNPVPDIRWRKVLEPMP-STAEIS-TSGAVLKIF 298
Qy    504 PASIR-PMKNITAIAGRDTYIHCRVIGYPYYSIKWYKNSNLLPFNHRQVAFENNGTLKLS 562

Db    299 NIQLE-DEGIYECEA-ENIRGKDKHQARIYVQAFPEWVEHINDTEVDIGSDLYWPCVATG 356
Qy    563 DVQKEVDEGEYTCNVLVQPQLSTSQSVHVTVKV-PPFIQPFEEPRFSIGQRVFIPCVVVS 621

Db    357 KPIP-TIRWLKNGYAYH-K-G------E----LRLYDV-ENAGMYQCIAENTYGAIYAN 403
Qy    622 GDLPITITWQKDGRPIPGSLGVTIDNIDFTSSLRISNLSLMHNGNYTCIARNEAAAVEHQ 681

Db    404 AELKILALAPTFEMNPMKKKILAAKGGRVIIECKPKAAPKPKFSW--SKGT---EWL-VN 457
Qy    682 SQL-IVRVPPKFVVQPRDQDGIYGKA--VILNCSAEGYPVPTIVWKFSKGAGVPQFQPIA 738

Db    458 -SSRILIWEDGSLEINNITRNDGGIYTCFAENNRGKANSTGT-LVITDPTRIILAPINAD 515
Qy    739 LNGRIQVLSNGSLLIKHVVEEDSGYYLCKVSNDVGADVSKSMYLTVKIPAMITSYPNTTL 798

Db    516 ITVGENATMQCAASFDPALDLTFVWSFNGYVIDFNKEN-IHYQRNFMLDSNGELLIRNAQ 574
Qy    799 ATQGQKKEMSCTAHGEKPI-IV-RWEKEDRIINPEMAPYLVSTKEVGEEVISTLQILPTV 856

Db    575 LKHAGRYTCTAQTIVDNSSASADLVVRGPPGPPGGLRIEDIRATSVALTWSRGSDNHSPI 634
Qy    857 REDSGFFSCHAINSYGEDRGIIQLTVQEPPDPPE-IEIKDVKARTITLRWTMGFDGNSPI 915

Db    635 SKYTIQTKTILSDDWKDAKTDPPIIEGNMEAARAVDLIPWMEYEFRVVATNTLGRGEPSI 694
Qy    916 TGYDIECKNK-SDSWDSAQRTKDV-SPQLNSATIIDIHPSSTYSIRMYAKNRIGKSEPSN 973

Db    695 PSNRIKTDGAAPNVAPSDVGGGGGRNRELTITWAPLSREYHYGNNFGYIVAFKPFD-GEE 753
Qy    974 ELT-ITADEAAPDGPPQEVHLEPISSQSIRVTWKAPKKHLQNGIIRGYQIGYREYSTGGN 1032

Db    754 WK-KVTVTNPDTGRYVHKDETMSPSTAFQVKVKAFNNKGDGPYSLLAVINSAQDAPSEAP 812
Qy   1033 FQFNIISVDTSGDSEVYTLDNLNKFTQYGLVVQACNRAGTGPSSQEIITTTLEDVPSYPP 1092

Db    813 TEVGVKVLSSSEISVHWE----HVLEKIVESYQIRYWAAHDKEEAANRVQVTSQEYSARL 868
Qy   1093 ENVQAIATSPESISISWSTLSKEALNGILQGFRVIYWANLMDGELGEIKNITTTQPSLEL 1152

Db    869 ENLLPDTQYFIEVGACNSAGCGPPSDMIEAFTKKAPPSQPPRIISSVRSGSRYIITWDHV 928
Qy   1153 DGLEKYTNYSIQVLAFTRAGDGVRSEQIFTRTKEDVPGPPAGVKAAAASASMVFVSW--L 1210

Db    929 VALSNESTVTGYKVLY-RPDGQHDGKL-YSTHKHSIEVP-IPRDGEYVVEVRAHSDGGDG 985
Qy   1211 PPLKLNGIIRKYTVFCSHPYPTVISEFEASPDSFSYRIPNLSRNRQYSVWVVAVTSAGRG 1270

Db    986 VVSQVKISGAP 996
Qy   1271 NSSEI-ITVEP 1280


RESULT    6
ID   US-08-714-901-6     STANDARD;     PRT;   1018 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 6, Application US/08714901
XX
CC    Sequence 6, Application US/08714901
CC    Patent No. 5739289
CC     GENERAL INFORMATION:
CC       APPLICANT:  Reid, Robert A.
CC       APPLICANT:  Hemperly, John J.
CC       TITLE OF INVENTION:  Human Cell Adhesion Molecule and Nucleic
CC       TITLE OF INVENTION:  Acid Sequences
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Richard J. Rodrick, Becton Dickinson
CC         ADDRESSEE:  and Company
CC         STREET:  One Becton Drive
CC         CITY:  Franklin Lakes
CC         STATE:  NJ
CC         COUNTRY:  USA
CC         ZIP:  07417
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/714,901
CC         FILING DATE:  17-SEP-1996
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/408,093
CC         FILING DATE:  21-MAR-1995
CC         APPLICATION NUMBER:  US/08/040,741
CC         FILING DATE:  26 MAR 1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Fugit, Donna R.
CC         REGISTRATION NUMBER:  32,135
CC         REFERENCE/DOCKET  UMBER:  P-2630
CC      INFORMATION FOR SEQ I  NO:  6:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1018 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       ORIGINAL SOURCE:
CC         ORGANISM:  Homo sapiens
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  45..94
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  138..191
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  243..290
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  332..371
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  416..464
CC       FEATURE:
CC         NAME/KEY:  Disulfide-bond
CC         LOCATION:  506..563
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  604..657
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
```

```
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  707..760
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  809..857
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
CC       FEATURE:
CC         NAME/KEY:  Domain
CC         LOCATION:  905..952
CC         OTHER INFORMATION:  /label= FLR
CC         OTHER INFORMATION:  /note= "conserved core of fibronectin type
CC         OTHER INFORMATION:  III-like repeat"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  188
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  238
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  318
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  437
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  453
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  474
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  501
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  571
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
CC       FEATURE:
CC         NAME/KEY:  Modified-site
CC         LOCATION:  913
CC         OTHER INFORMATION:  /label= ASN-glycos
CC         OTHER INFORMATION:  /note= "potential site of ASN-linked
CC         OTHER INFORMATION:  glycosylation"
SQ   SEQUENCE   1018 AA; 113333 MW; 5493875 CN;
```

```
  Query Match           5.6%;  Score 629;  DB 1;  Length 1018;
  Best Local Similarity  25.0%;  Pred. No. 4.68e-33;
  Matches  198;  Conservative  196;  Mismatches 348;  Indels  49;  Gaps  33;

Db    241 PADIVVQFKDVYALMGQNVTLECFALGNPVPDIRWRKVLEPMP-STAEIS-TSGAVLKIF 298
Qy    504 PASIR-PMKNITAIAGRDTYIHCRVIGYPYYSIKWYKNSNLLPFNHRQVAFENNGTLKLS 562

Db    299 NIQLE-DEGIYECEA-ENIRGKDKHQARIYVQAFPEWVEHINDTEVDIGSDLYWPCVATG 356
Qy    563 DVQKEVDEGEYTCNVLVQPQLSTSQSVHVTVKV-PPFIQPFEFPRFSIGQRVFIPCVVVS 621

Db    357 KPIP-TIRWLKNGYAYH-K-G------E----LRLYDVTFENAGMYQCIAENTYGAIYAN 403
Qy    622 GDLPITITWQKDGRPIPGSLGVTIDNIDFTSSLRISNLSLMHNGNYTCIARNEAAAVEHQ 681

Db    404 AELKILALAPTFEMNPMKKKILAAKGGRVIIECKPKAAPKPKFSW--SKGT---EWL-VN 457
Qy    682 SQL-IVRVPPKFVVQPRDQDGIYGKA--VILNCSAEGYPVPTIVWKFSKGAGVPQFQPIA 738

Db    458 -SSRILIWEDGSLEINNITRNDGGIYTCFAENNRGKANSTGT-LVITDPTRIILAPINAD 515
Qy    739 LNGRIQVLSNGSLLIKHVVEEDSGYYLCKVSNDVGADVSKSMYLTVKIPAMITSYPNTTL 798

Db    516 ITVGENATMQCAASFDPALDLTFVWSFNGYVIDFNKEN-IHYQRNFMLDSNGELLIRNAQ 574
Qy    799 ATQGQKKEMSCTAHGEKPI-IV-RWEKEDRIINPEMARYLVSTKEVGEEVISTLQILPTV 856

Db    575 LKHAGRYTCTAQTIVDNSSASADLVVRGPPGPPGGLRIEDIRATSVALTWSRGSDNHSPI 634
Qy    857 REDSGFFSCHAINSYGEDRGIIQLTVQEPPDPPE-IEIKDVKARTITLRWTMGFDGNSPI 915

Db    635 SKYTIQTKTILSDDWKDAKTDPPIIEGNMEAARAVDLIPWMEYEFRVVATNTLGRGEPSI 694
Qy    916 TGYDIECKNK-SDSWDSAQRTKDV-SPQLNSATIIDIHPSSTYSIRMYAKNRIGKSEPSN 973

Db    695 PSNRIKTDGAAPNVAPSDVGGGGGRNRELTITWAPLSREYHYGNNFGYIVAFKPFD-GEE 753
Qy    974 ELT-ITADEAAPDGPPQEVHLEPISSQSIRVTWKAPKKHLQNGIIRGYQIGYREYSTGGN 1032

Db    754 WK-KVTVTNPDTGRYVHKDETMSPSTAFQVKVKAFNNKGDGPYSLLAVINSAQDAPSEAP 812
Qy   1033 FQFNIISVDTSGDSEVYTLDNLNKFTQYGLVVQACNRAGTGPSSQEIITTTLEDVPSYPP 1092

Db    813 TEVGVKVLSSSEISVHWE----HVLEKIVESYQIRYWAAHDKEEAANRVQVTSQEYSARL 868
Qy   1093 ENVQAIATSPESISISWSTLSKEALNGILQGFRVIYWANLMDGELGEIKNITTTQPSLEL 1152

Db    869 ENLLPDTQYFIEVGACNSAGCGPPSDMIEAFTKKAPPSQPPRIISSVRSGSRYIITWDHV 928
Qy   1153 DGLEKYTNYSIQVLAFTRAGDGVRSEQIFTRTKEDVPGPPAGVKAAAASASMVFVSW--L 1210

Db    929 VALSNESTVTGYKVLY-RPDGQHDGKL-YSTHKHSIEVP-IPRDGEYVVEVRAHSDGGDG 985
Qy   1211 PPLKLNGIIRKYTVFCSHPYPTVISEFEASPDSFSYRIPNLSRNRQYSVWVVAVTSAGRG 1270

Db    986 VVSQVKISGAP 996
Qy   1271 NSSEI-ITVEP 1280


RESULT    7
ID   US-08-452-052-2     STANDARD;     PRT;   1018 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application US/08452052
XX
CC    Sequence 2, Application US/08452052
CC    Patent No. 5766922
```

```
CC     GENERAL INFORMATION:
CC       APPLICANT:  Peles, Elior
CC       TITLE OF INVENTION:  FUNCTIONAL LIGANDS FOR THE AXONAL CELL
CC       TITLE OF INVENTION:  RECOGNITION MOLECULE CONTACTIN
CC       NUMBER OF SEQUENCES:  2
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pennie & Edmonds
CC         STREET:  1155 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  U.S.A.
CC         ZIP:  10036-2711
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/452,052
CC         FILING DATE:  26-MAY-1995
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Coruzzi, Laura A.
CC         REGISTRATION NUMBER:  30,742
CC         REFERENCE/DOCKET NUMBER:  7683-075
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 790-9090
CC         TELEFAX:  (212) 869-8864/9741
CC         TELEX:  66141 PENNIE
CC      INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1018 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:
CC         TOPOLOGY:  unknown
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   1018 AA; 113102 MW; 5593274 CN;

  Query Match          5.4%;  Score 606;  DB 1;  Length 1018;
  Best Local Similarity  24.9%;  Pred. No. 1.86e-31;
  Matches  200;  Conservative  205;  Mismatches 344;  Indels  54;  Gaps  37;

Db    240 PADIVVQFKDIYTMMGQNVTLECFALGNPVPDIRWRKVLEPMP-TTAEIS-TSGAVLKIF 297
Qy    504 PASIR-PMKNITAIAGRDTYIHCRVIGYPYYSIKWYKNSNLLPFNHRQVAFENNGTLKLS 562

Db    298 NIQLE-DEGLYECEA-ENIRGKDKHQARIYVQAFPEWVEHINDTEVDIGSDLYWPCVATG 355
Qy    563 DVQKEVDEGEYTCNVLVQPQLSTSQSVHVTVKV-PPFIQPFEFPRFSIGQRVFIPCVVVS 621

Db    356 KPIP-TIRWLKNGYAYH-K-G------E----LRLYDVTFENAGMYQCIAENAYGTIYAN 402
Qy    622 GDLPITITWQKDGRPIPGSLGVTIDNIDFTSSLRISNLSLMHNGNYTCIARNEAAAVEHQ 681

Db    403 AELKILALAPTFEMNPMKKKILAAKGGRVIIECKPKAAPKPKFSW--SKGT---EWL-VN 456
Qy    682 SQL-IVRVPPKFVVQPRDQDGIYGKA--VILNCSAEGYPVPTIVWKFSKGAGVPQFQPIA 738

Db    457 -SSRILIWEDGSLEINNITRNDGGIYTCFAENNRGKANSTGT-LVITNPTRIILAPINAD 514
Qy    739 LNGRIQVLSNGSLLIKHVVEEDSGYYLCKVSNDVGADVSKSMYLTVKIPAMITSYPNTTL 798

Db    515 ITVGENATMQCAASFDPSLDLTFVWSFNGYVIDFNKEITHIHYQRNFMLDANGELLIRNA 574
Qy    799 ATQGQKKEMSCTAHGEKPI-IV-RWEKEDRIINPEMARYLVS-TKEVGEEVISTLQILPT 855

Db    575 QLKHAGRYTCTAQTIVDNSSASADLVVRGPPGPPGGLRIEDIRATSVALTWSRGSDNHSP 634
Qy    856 VREDSGFFSCHAINSYGEDRGIIQLTVQEPPDPPE-IEIKDVKARTITLRWTMGFDGNSP 914

Db    635 ISKYTIQTKTILSDDWKDAKTDPPIIEGNMESAKAVDLIPWMEYEFRVVATNTLGTGEPS 694
Qy    915 ITGYDIECKNK-SDSWDSAQRTKDV-SPQLNSATIIDIHPSSTYSIRMYAKNRIGKSEPS 972

Db    695 IPSNRIKTDGAAPNVAPSDVGGGGGTNRELTITWAPLSREYHYGNNFGYIVAFKPFD--G 752
Qy    973 NELT-ITADEAAPDGPPQEVHLEPISSQSIRVTWKAPKKHLQNGIIRGYQIGYREYSTGG 1031

Db    753 E-EWKKVTV-TNPDTGRYVHKETMTPSTAFQVKVKAFNNKGDGPYSLIAVINSAQDAPSE 810
Qy   1032 NFQFNIISVDTSGDSEVYTL-DNLNKFTQYGLVVQACNRAGTGPSSQEIITTTLEDVPSY 1090

Db    811 APTEVGVKVLSSSEISVHW----KHVLEKIVESYQIRYAGHDKEAAAHRVQ-VTSQEYSA 865
Qy   1091 PPENVQAIATSPESISISWSTLSKEALNGILQGFRVIYWANLMDGELGEIKNITTTQPSL 1150

Db    866 RLENLLPDTQYFIEVGACNSAGCGPSSDVIETFTRKAPPSQPPRIISSVRSGSRYIITWD 925
Qy   1151 ELDGLEKYTNYSIQVLAFTRAGDGVRSEQIFTRTKEDVPGPPAGVKAAAASASMVFVSW- 1209

Db    926 HVVALSNESTVTGYKILY-RPDGQHDGKL-FSTHKHSIEVP-IPRDGEYVVEVRAHSDGG 982
Qy   1210 -LPPLKLNGIIRKYTVFCSHPYPTVISEFEASPDSFSYRIPNLSRNRQYSVWVVAVTSAG 1268

Db    983 DGVVSQVKISGVSTLSSGLLSLL 1005
Qy   1269 RGNSSEI-IT-VEPLAKAPARIL 1289


RESULT    8
ID   US-08-348-006B-5     STANDARD;     PRT;   1911 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 5, Application US/08348006B
XX
CC    Sequence 5, Application US/08348006B
CC    Patent No. 5658756
CC     GENERAL INFORMATION:
CC       APPLICANT:  RODAN, GIDEON A.
CC       APPLICANT:  SCHMIDT, AZRIEL
CC       APPLICANT:  RUTLEDGE, SU JANE
CC       TITLE OF INVENTION:  CDNA ENCODING A NOVEL HUMAN PROTEIN
CC       TITLE OF INVENTION:  TYROSINE PHOSPHATASE
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  J. MARK HAND
CC         STREET:  126 E. LINCOLN AVE., P.O. BOX 2000
CC         CITY:  RAHWAY
CC         STATE:  NEW JERSEY
CC         COUNTRY:  USA
CC         ZIP:  07065-0900
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/348,006B
CC         FILING DATE:
CC         CLASSIFICATION:  514
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/122,032
CC         FILING DATE:  14-SEP-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  HAND, J., MARK
CC         REGISTRATION NUMBER:  36,545
CC         REFERENCE/DOCKET NUMBER:  18992IA
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  908-594-3905
CC         TELEFAX:  908-594-4720
CC      INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1911 amino acids
```

```
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   1911 AA; 213063 MW; 19256861 CN;

  Query Match            5.1%;  Score 569;  DB 1;  Length 1911;
  Best Local Similarity  25.6%;  Pred. No. 6.80e-29;
  Matches   230;  Conservative  217;  Mismatches 379;  Indels  74;  Gaps  60;

Db      33 PRFIKEPK-DQIGVSGRVASFV-CQA-TGDPKPRVTWNKKGKKVNSQRFETIEFDESAGA 89
Qy     596 PPFIQPFEFPRFSIGQRV--FIPCVVVSGDLPITITWQKDGRPIPGSLGVTIDNIDFTSS 653

Db      90 VLRIQPLRTPRDENVYECVAQNSVGEITVHAKLTVLREDQLPSGFPNIDMGPQLKVVERT 149
Qy     654 -LRISNLSLMHNGN-YTCIARNEAAAVEHQSQLIV-R---VPPKFV-VQPRDQ-DGIY-G 704

Db     150 RTATMLCAASGNPDPEITW-F-KDF-LP-VDPSASNGRIKQLRSGALQIESSEETDQGKY 205
Qy     705 KAVILNCSAEGYPVPTIVWKFSKGAGVPQFQPIALNGRIQVLSNGSLLIKHVVEEDSGYY 764

Db     206 ECVATNSAGVRYSSPANLYVRVRRVAPRFSILPMSHEIMPGGNVNITCVAVGSPMPYVKW 265
Qy     765 LCKVSNDVGADVSK-S-MYLTVK-IPAMITSYPNTTLATQGQKKEMSCTAHGEKPIIVRW 821

Db     266 -MQG--AE-DLTP--EDDMPVGRNV---LE-LTDVKDSANYHPCVAMSSLGVIEAVAQIT 315
Qy     822 EKEDRIINPEMARYLVSTKEVGEEVISTLQILPTVREDSGFFSCHAINSYGEDRGIIQLT 881

Db     316 VKSLPKAPGTPMVTENTATSITIITWDSG-NPD-PVSYYVIEYKSKSQD-G[illegible]IKEDITT 372
Qy     882 VQEPPDPPEIE-IKDVKARTITLRWTMGFDGNSPITGYDIECKNKSDSWDSAQRTKDVSP 940

Db     373 --TRYSIGGLSPNSEYEIWVSAVNSIGQGPPSESVVTRTGEQAPARPPRNVQARMLSATT 430
Qy     941 QLNSATIIDIHPSSTYSIRMYAKNRIGKSEPSNELTITADEAAPDGPPQEVHLEPISSQS 1000

Db     431 MIVQWEEPV-E-PNGLIRGYRV-Y--YTMEPEHPVGNWQ[illegible]NVDDSLLTTVGSLLEDETY 485
Qy    1001 IRVTWKAPKKHLQNGIIRGYQIGYREYSTGGNFQFNIISVDTSGDSEVYTLDNLNKFTQY 1060

Db     486 TVRVLAFTSVGDGPLSDPIQVKTQQGVPG-QPMNLRAEARSETSITLSWSP-PRQ--ESI 541
Qy    1061 GLVVQACNRAGTGPSSQEIITTTLEDVPSYPPENVQAIATSPESISISWSTLSKEALNGI 1120

Db     542 IK-YELLFREGDHGREVG--RTFDPT-TSYVVEDLKPNTEYAFRLAARSPQGLGAFTPVV 597
Qy    1121 LQGFRVIYWANLMDGELGEIKNITTTQPSLELDGLEKYTNYSIQVLAFTRAGDGVRSEQI 1180

Db     598 RQRTLQSKPSAPPQDVKCVSVRSTAILVSWRPPPPETHNGALVGYSVRYRPLGSEDPEPK 657
Qy    1181 FTRTKEDVPG-PPAGVKAAAASASMVFVSW-LPPLKL-NGIIRKYTV-FCSHPYPTVIS- 1235

Db     658 EVNGIPPTTTQILLEALEKWTQYRITTVAHTEVGPGPESSPVVVRTDEDVPSAPPRKVEA 717
Qy    1236 EFEA-SPDSFSYRIPNLSRNRQYSVWVVAVTSAGRG-NSSEII--TVEPLAKAPAR-I-L 1289

Db     718 EALNATAIRVLW-RSPA-PGRQHGQIRGYQVHYVR-MEGAEARGPPRIKDVMLADAQEMV 774
Qy    1290 -TFSGT-VTTPWMKDIVLPCKAVGDPSP-AVKWMKDSNGTPSLVTIDGRRSIFSNG-SFI 1345

Db     775 ITNLQPE-TAYSITVAAYTMKGDGARSKPKVVVTKGAVLGRPTLSVQQTPEGSLLARWEP 833
Qy    1346 IRTVKAEDSGYYSCIANNNWGSDEIILNLQVQVP--P--DQPRLTVSKTTSSSITLSW-L 1400

Db     834 PAGTAEDQVLGYRLQFGREDSTPLATLEFPPSEDRYTASGVHKGATYVFRLAARSPGGLG 893
Qy    1401 PGDNGGSSIRGYILQYSEDNSEQWGSFPISPSERSYRLENLKCGTWYKFTLTAQNGVGPG 1460

RESULT    9
ID   US-08-800-825A-5     STANDARD;     PRT;   1911 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 5, Application US/08800825A
XX
CC    Sequence 5, Application US/08800825A
CC    Patent No. 5866397
CC     GENERAL INFORMATION:
CC       APPLICANT:  RODAN, GIDEON A.
CC       APPLICANT:  SCHMIDT, AZRIEL
CC       APPLICANT:  RUTLEDGE, SU JANE
CC       TITLE OF INVENTION:  CDNA ENCODING A NOVEL HUMAN PROTEIN
CC       TITLE OF INVENTION:  TYROSINE PHOSPHATASE
CC       NUMBER OF SEQUENCES:  7
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  J. MARK HAND - MERCK & CO., INC.
CC         STREET:  126 E. LINCOLN AVE., P.O. BOX 2000
CC         CITY:  RAHWAY
CC         STATE:  NEW JERSEY
CC         COUNTRY:  USA
CC         ZIP:  07065-0900
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/800,825A
CC         FILING DATE:  14-FEB-1997
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  HAND, J. MARK
CC         REGISTRATION NUMBER:  36,545
CC         REFERENCE/DOCKET NUMBER:  18992DA
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  732-594-3905
CC         TELEFAX:  732-594-4720
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1911 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   1911 AA; 213063 MW; 19256861 CN;

  Query Match            5.1%;  Score 569;  DB 2;  Length 1911;
  Best Local Similarity  25.6%;  Pred. No. 6.80e-29;
  Matches   230;  Conservative  217;  Mismatches 379;  Indels  74;  Gaps  60;

Db      33 PRFIKEPK-DQIGVSGRVASFV-CQA-TGDPKPRVTWNKKGKKVNSQRFETIEFDESAGA 89
Qy     596 PPFIQPFEFPRFSIGQRV--FIPCVVVSGDLPITITWQKDGRPIPGSLGVTIDNIDFTSS 653

Db      90 VLRIQPLRTPRDENVYECVAQNSVGEITVHAKLTVLREDQLPSGFPNIDMGPQLKVVERT 149
Qy     654 -LRISNLSLMHNGN-YTCIARNEAAAVEHQSQLIV-R---VPPKFV-VQPRDQ-DGIY-G 704

Db     150 RTATMLCAASGNPDPEITW-F-KDF-LP-VDPSASNGRIKQLRSGALQIESSEETDQGKY 205
Qy     705 KAVILNCSAEGYPVPTIVWKFSKGAGVPQFQPIALNGRIQVLSNGSLLIKHVVEEDSGYY 764

Db     206 ECVATNSAGVRYSSPANLYVRVRRVAPRFSILPMSHEIMPGGNVNITCVAVGSPMPYVKW 265
Qy     765 LCKVSNDVGADVSK-S-MYLTVK-IPAMITSYPNTTLATQGQKKEMSCTAHGEKPIIVRW 821

Db     266 -MQG--AE-DLTP--EDDMPVGRNV---LE-LTDVKDSANYHPCVAMSSLGVIEAVAQIT 315
Qy     822 EKEDRIINPEMARYLVSTKEVGEEVISTLQILPTVREDSGFFSCHAINSYGEDRGIIQLT 881

Db     316 VKSLPKAPGTPMVTENTATSITIITWDSG-NPD-PVSYYVIEYKSKSQD-GPYQIKEDITT 372
```

```
Qy    882 VQEPPDPPEIE-IKDVKARTITLRWTMGFDGNSPITGYDIECKNKSDSWDSAQRTKDVSP 940

Db    373 --TRYSIGGLSPNSEYEIWVSAVNSIGQGPPSESVVTRTGEQAPARPPRNVQARMLSATT 430

Qy    941 QLNSATIIDIHPSSTYSIRMYAKNRIGKSEPSNELTITADEAAPDGPPQEVHLEPISSQS 1000

Db    431 MIVQWEEPV-E-PNGLIRGYRV-Y--YTMEPEHPVGNWQKHNVDDSLLTTVGSLLEDETY 485

Qy   1001 IRVTWKAPKKHLQNGIIRGYQIGYREYSTGGNFQFNIISVDTSGDSEVYTLDNLNKFTQY 1060

Db    486 TVRVLAFTSVGDGPLSDPIQVKTQQGVPG-QPMNLRAEARSETSITLSWSP-PRQ--ESI 541

Qy   1061 GLVVQACNRAGTGPSSQEIITTTLEDVPSYPPENVQAIATSPESISISWSTLSKEALNGI 1120

Db    542 IK-YELLFREGDHGREVG--RTFDPT-TSYVVEDLKPNTEYAFRLAARSPQGLGAFTPVV 597

Qy   1121 LQGFRVIYWANLMDGELGEIKNITTTQPSLELDGLEKYTNYSIQVLAFTRAGDGVRSEQI 1180

Db    598 RQRTLQSKPSAPPQDVKCVSVRSTAILVSWRPPPPETHNGALVGYSVRYRPLGSEDPEPK 657

Qy   1181 FTRTKEDVPG-PPAGVKAAAASASMVFVSW-LPPLKL-NGIIRKYTV-FCSHPYPTVIS- 1235

Db    658 EVNGIPPTTTQILLEALEKWTQYRITTVAHTEVGPGPESSPVVVRTDEDVPSAPPRKVEA 717

Qy   1236 EFEA-SPLSFSYRIPNLSRNRQYSVWVVAVTSAGRG-NSSEII--TVEPLAKAPAR-I-L 1289

Db    718 EALNATAIRVLW-RSPA-PGRQHGQIRGYQVHYVR-MEGAEARGPPRIKDVMLADAQEMV 774

Qy   1290 -TFSGT-VTTPWMKDIVLPCKAVGDPSP-AVKWMKDSNGTPSLVTIDGRRSIFSNG-SFI 1345

Db    775 ITNLQPE-TAYSITVAAYTMKGDGARSKPKVVVTKGAVLGRPTLSVQQTPEGSLLARWEP 833

Qy   1346 IRTVKAEDSGYYSCIANNNWGSDEIILNLQVQVP--P--DQPRLTVSKTTSSSITLSW-L 1400

Db    834 PAGTAEDQVLGYRLQFGREDSTPLATLEFPPSEDRYTASIVHKGATYVFRLAARSPGGLG 893

Qy   1401 PGDNGGSSIRGYILQYSEDNSEQWGSFPISPSERSYRLENLKCGTWYKFTLTAQNGVGPG 1460


RESULT  10
ID   PCT-US94-10166-5     STANDARD;     PRT;   1911 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 5, Application PC/TUS9410166
XX
CC    Sequence 5, Application PC/TUS9410166
CC     GENERAL INFORMATION:
CC       APPLICANT:  RODAN, GIDEON A
CC       APPLICANT:  SCHMIDT, AZRIEL
CC       APPLICANT:  RUTLEDGE, SU JANE
CC       TITLE OF INVENTION:  CDNA ENCODING A NOVEL HUMAN PROTEIN
CC       TITLE OF INVENTION:  TYROSINE PHOSPHATASE
CC       NUMBER OF SEQUENCES:  6
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  JOHN W. WALLEN III
CC         STREET:  P.O. BOX 2000, 126 E. LINCOLN AVE.
CC         CITY:  RAHWAY
CC         STATE:  NJ
CC         COUNTRY:  USA
CC         ZIP:  07065
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/10166
CC         FILING DATE:  09-SEPT-1994
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/122,032
CC         FILING DATE:  14-SEP-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  WALLEN, JOHN W III
CC         REGISTRATION NUMBER:  35403
CC         REFERENCE/DOCKET NUMBER:  18992
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  908-594-3905
CC         TELEFAX:  908-594-4720
CC         TELEX:  138825
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1911 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   1911 AA; 213063 MW; 19256861 CN;

  Query Match             5.1%;  Score 569;  DB 3;  Length 1911;
  Best Local Similarity  25.6%;  Pred. No. 6.80e-29;
  Matches   230;  Conservative  217;  Mismatches 379;  Indels  74;  Gaps  60;

Db     33 PRFIKEPK-DQIGVSGRVASFV-CQA-TGDPKPRVTWNKKGKKVNSQRFETIEFDESAGA 89

Qy    596 PPFIQPFEFPRFSIGQRV--FIPCVVVSGDLPITITWQKDGRPIPGSLGVTIDNIDFTSS 653

Db     90 VLRIQPLRTPRDENVYECVAQNSVGEITVHAKLTVLREDQLPSGFPNIDMGPQLKVVERT 149

Qy    654 -LRISNLSLMHNGN-YTCIARNEAAAVEHQSQLIV-R---VPPKFV-VQPRDQ-DGIY-G 704

Db    150 RTATMLCAASGNPDPEITW-F-KDF-LP-VDPSASNGRIKQLRSGALQIESSEETDQGKY 205

Qy    705 KAVILNCSAEGYPVPTIVWKFSKGAGVPQFQPIALNGRIQVLSNGSLLIKHVVEEDSGYY 764

Db    206 ECVATNSAGVRYSSPANLYVRVRRVAPRFSILPMSHEIMPGGNVNITCVAVGSPMPYVKW 265

Qy    765 LCKVSNDVGADVSK-S-MYLTVK-IPAMITSYPNTTLATQGQKKEMSCTAHGEKPIIVRW 821

Db    266 -MQG--AE-DLTP--EDLMPVGRNV---LE-LTDVKDSANYHPCVAMSSLGVIEAVAQIT 315

Qy    822 EKEDRIINPEMARYLVSTKEVGFEVISTLQILPTVREDSGFFSCHAINSYGEDRGIIQLT 881

Db    316 VKSLPKAPGTPMVTENTATSITITWDSG-NPD-PVSYYVIEYKSKSQD-GPYQIKEDITT 372

Qy    882 VQEPPDPPEIE-IKDVKARTITLRWTMGFDGNSPITGYDIECKNKSDSWDSAQRTKDVSP 940

Db    373 --TRYSIGGLSPNSEYEIWVSAVNSIGQGPPSESVVTRTGEQAPARPPRNVQARMLSATT 430

Qy    941 QLNSATIIDIHPSSTYSIRMYAKNRIGKSEPSNELTITADEAAPDGPPQEVHLEPISSQS 1000

Db    431 MIVQWEEPV-E-PNGLIRGYRV-Y--YTMEPEHPVGNWQKHNVDDSLLTTVGSLLEDETY 485

Qy   1001 IRVTWKAPKKHLQNGIIRGYQIGYREYSTGGNFQFNIISVDTSGDSEVYTLDNLNKFTQY 1060

Db    486 TVRVLAFTSVGDGPLSDPIQVKTQQGVPG-QPMNLRAEARSETSITLSWSP-PRQ--ESI 541

Qy   1061 GLVVQACNRAGTGPSSQEIITTTLEDVPSYPPENVQAIATSPESISISWSTLSKEALNGI 1120

Db    542 IK-YELLFREGDHGREVG--RTFDPT-TSYVVEDLKPNTEYAFRLAARSPQGLGAFTPVV 597

Qy   1121 LQGFRVIYWANLMDGELGEIKNITTTQPSLELDGLEKYTNYSIQVLAFTRAGDGVRSEQI 1180

Db    598 RQRTLQSKPSAPPQDVKCVSVRSTAILVSWRPPPPETHNGALVGYSVRYRPLGSEDPEPK 657

Qy   1181 FTRTKEDVPG-PPAGVKAAAASASMVFVSW-LPPLKL-NGIIRKYTV-FCSHPYPTVIS- 1235

Db    658 EVNGIPPTTTQILLEALEKWTQYRITTVAHTEVGPGPESSPVVVRTDEDVPSAPPRKVEA 717

Qy   1236 EFEA-SPDSFSYRIPNLSRNRQYSVWVVAVTSAGRG-NSSEII--TVEPLAKAPAR-I-L 1289

Db    718 EALNATAIRVLW-RSPA-PGRQHGQIRGYQVHYVR-MEGAEARGPPRIKDVMLADAQEMV 774
```

```
          ::::| :   | :    | :  |:     |  ::  :|: :      :  :::::   :
Qy   1290 -TFSGT-VTTPWMKDIVLPCKAVGDPSP-AVKWMKDSNGTPSLVTIDGRRSIFSNG-SFI 1345

Db    775 ITNLQPE-TAYSITVAAYTMKGDGARSKPKVVVTKGAVLGRPTLSVQQTPEGSLLARWEP 833
          |  : :| ::|   :|     :|    :  | |   :   :| |:|  |  :|:   |
Qy   1346 IRTVKAEDSGYYSCIANNNWGSDEIILNLQVQVP--P--DQPRLTVSKTTSSSITLSW-L 1400

Db    834 PAGTAEDQVLGYRLQFGREDSTPLATLEFPPSEDRYTASGVHKGATYVFRLAARSPGGLG 893
          |:  :   : || ||:: ::|   ::: ::||| |    :  |: | | |:|::  | |
Qy   1401 PGDNGGSSIRGYILQYSEDNSEQWGSFPISPSERSYRLENLKCGTWYKFTLTAQNGVGPG 1460


RESULT   11
ID   US-08-716-679-3     STANDARD;     PRT;   1501 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 3, Application US/08716679
XX
CC    Sequence 3, Application US/08716679
CC    Patent No. 5846800
CC     GENERAL INFORMATION:
CC       APPLICANT:  Schlessinger, Joseph
CC       APPLICANT:  Yan, Hai
CC       TITLE OF INVENTION:  NOVEL RECEPTOR-TYPE PROTEIN
CC       TITLE OF INVENTION:  PHOSPHOTYROSINE PHOSPHATASE-SIGMA
CC       NUMBER OF SEQUENCES:  12
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pennie & Edmonds
CC         STREET:  1155 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  U.S.A.
CC         ZIP:  10036-2711
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/716,679
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/130,570
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Misrock, S. Leslie
CC         REGISTRATION NUMBER:  18,872
CC         REFERENCE/DOCKET NUMBER:  7683-043
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  212-790-9090
CC         TELEFAX:  212-869-8864/9741
CC         TELEX:  66141 PENNIE
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1501 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   1501 AA; 168336 MW; 11919615 CN;

  Query Match             4.3%;  Score 486;  DB 2;  Length 1501;
  Best Local Similarity  26.8%;  Pred. No. 3.52e-23;
  Matches   163;  Conservative  147;  Mismatches 249;  Indels  50;  Gaps  40;

Db     52 FV-CQA-TGDPKPRVTWNKKGKKVNSQRFETIDFDESSGAVLRIQPLRTPRDENVYECVA 109
          |: |   :||    :|| | |: : :    |||  : ::: |||  |   :: | | |:|
Qy    614 FIPCVVVSGDLPITITWQKDGRPIPGSLGVTIDNIDFTSS-LRISNLSLMHNGN-YTCIA 671

Db    110 QNSVGEITVHAKLTVLREDQLPPGFPNIDMGPQLKVVERTRTATMLCAASGNPDPEITW- 168
          :|  : :  :: | | |   :|| |  ::   |   :   ::  : |:| | | | | |
Qy    672 RNEAAAVEHQSQLIV-R---VPPKFV-VQPRDQ-DGIY-GKAVILNCSAEGYPVPTIVWK 724

Db    169 F-KDF-LP-VDPSASNGRIKQLRSGALQIESSEETDQGKYECVATNSAGVRYSSPANLYV 225
          | |   :|  :| | |||| | :|:| |    | | | | |  :|  |   | : :|:
Qy    725 FSKGAGVPQFQPIALNGRIQVLSNGSLLIKHVVEEDSGYYLCKVSNDVGADVSK-S-MYL 782

Db    226 RVRRVAPRFSILPMSHEIMPGGNVNITCVAVGSPMPYVKW-MQG--AE-DLTP--EDDMP 279
          |: ::: ::  | :     | : :::| | |     |:| ::    : :::      :
Qy    783 TVK-IPAMITSYPNTTLATQGQKKEMSCIAHGEKPIIVRWEKEDRIINPEMARYLVSTKE 841

Db    280 VGRNV---LE-LTDVKDSANY-TCVAMSSLGVIEAVAQITVKSLPKAPGTPVVTENTATS 334
          || :|   |: |  |:: : : :| |::| |   :: |:||   | :|   :  :  | :
Qy    842 VGEEVISTLQILPTVREDSGFFSCHAINSYGEDRGIIQLTVQEPPDPPEIEI-KDVKART 900

Db    335 ITVTWDSG-NPD-PVSYYVIEYKSKSQD-GPYQIKEDITT--TRYSIGGLSPNSEYEIWV 389
          ||: |  | : : |:: | || |:||:   : |   |::      :|  : |:| | |::
Qy    901 ITLRWTMGFDGNSPITGYDIECKNKSDSWDSAQRTKDVSPQLNSATIIDIHPSSTYSIRM 960

Db    390 SAVNSIGQGPPSESVVTRTGEQAPASAPRNVQARMLSATTMIVQWEEPV-E-PNGLIRGY 447
          | | || : ||: :   : | || ::|::|:   :|: :: | |  |     ||:||||
Qy    961 YAKNRIGKSEPSNELTITADEAAPDGPPQEVHLEPISSQSIRVTWKAPKKHLQNGIIRGY 1020

Db    448 RV-Y--YTMEPEHPVGNWQKHNVDDSLLTTVGSLLEDETYTVRVLAFTSVGDGPLSDPIQ 504
          :: |  |:   :            || : |: :|     | : | |    | || |: |
Qy   1021 QIGYREYSTGGNFQFNIISVDTSGDSEVYTLDNLNKFTQYGLVVQACNRAGTGPSSQEII 1080

Db    505 VKTQQGVPG-QPMNLRAEAKSETSIGLSWSA-PRQ--ESVIK-YELLFREGDRGREVG-- 557
          | : ||:  | |::| | |  ||::|||: :::  ::::  : :::      |:|
Qy   1081 TTTLEDVPSYPPENVQAIATSPESISISWSTLSKEALNGILQGFRVIYWANLMDGELGEI 1140

Db    558 RTFDPT-TAFVVEDLKPNTEYAFRLAARSPQGLGAFTAVVRQRTLQAISPKNFKVKMIM- 615
          : :  |  :: :: |   |:|:::: | :  | |  :  :  || : ::     ||
Qy   1141 KNITTTQPSLELDGLEKYTNYSIQVLAFTRAGDGVRSEQIFTRTKEDVPGPPAGVKAAAA 1200

Db    616 KTS-VLLSW 623
            :| |::||
Qy   1201 SASMVFVSW 1209


RESULT   12
ID   US-08-447-464-3     STANDARD;     PRT;   1501 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 3, Application US/08447464
XX
CC    Sequence 3, Application US/08447464
CC    Patent No. 5840842
CC     GENERAL INFORMATION:
CC       APPLICANT:  Schlessinger, Joseph
CC       APPLICANT:  Yan, Hai
CC       TITLE OF INVENTION:  NOVEL RECEPTOR-TYPE PROTEIN
CC       TITLE OF INVENTION:  PHOSPHOTYROSINE PHOSPHATASE-SIGMA
CC       NUMBER OF SEQUENCES:  12
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Pennie & Edmonds
CC         STREET:  1155 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  U.S.A.
CC         ZIP:  10036-2711
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/447,464
CC         FILING DATE:  24-MAY-1995
CC         CLASSIFICATION:  435
```

```
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/130,570
CC         FILING DATE:  01-OCT-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Misrock, S. Leslie
CC         REGISTRATION NUMBER:  18,872
CC         REFERENCE/DOCKET NUMBER:  7683-043
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  212-790-9090
CC         TELEFAX:  212-869-8864/9741
CC         TELEX:  66141 PENNIE
CC     INFORMATION FOR SEQ ID NO:  3:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  1501 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   1501 AA; 168336 MW; 11919615 CN;

  Query Match             4.3%;  Score 486;  DB 2:  Length 1501;
  Best Local Similarity  26.8%;  Pred. No. 3.52e-23;
  Matches   163;  Conservative  147;  Mismatches 249;  Indels  50;  Gaps  40;

Db     52 FV-CQA-TGDPKPRVTWNKKGKKVNSQRFETIDFDESSGAVLRIQPLRTPRDENVYECVA 109
          |: |   :||    :|| | |: : :    |||  : ::: |||  |   :: | | |:|
Qy    614 FIPCVVVSGDLPITITWQKDGRPIPGSLGVTIDNIDFTS-LRISNLSLMHNGN-YTCIA 671

Db    110 QNSVGEITVHAKLTVLREDQLPPGFPNIDMGPQLKVVERTRTATMLCAASGNPDPEITW- 168
          :|  : :  :: | | |   :|| |  ::   |   :   ::  : |:| | | | | |
Qy    672 RNEAAAVEHQSQLIV-R---VPPKFV-VQPRDQ-DGIY-GKAVILNCSAEGYPVPTIVWK 724

Db    169 F-KDF-LP-VDPSASNGRIKQLRSGALQIESSEETDQGKYECVATNSAGVRYSSPANLYV 225
          | |   :|  :| | ||||  | :|:| |    | | | | |  :|  |   |  : :|:
Qy    725 FSKGAGVPQFQPIALNGRIQVLSNGSLLIKHVVEEDSGYYLCKVSNDVGADVSK-S-MYL 782

Db    226 RVRRVAPRFSILPMSHEIMPGGNVNITCVAVGSPMPYVKW-MQG--AE-DLTP--EDDMP 279
           |: ::: ::  | :     | : :::| | |     |:| ::    : :::      :
Qy    783 TVK-IPAMITSYPNTTLATQGQKKEMSCTAHGEKPIIVRWEKEDRIINPEMARYLVSTKE 841

Db    280 VGRNV---LE-LTDVKDSANY-TCVAMSSLGVIEAVAQITVKSLPKAPGTPVVTENTATS 334
          || :|   |: | |:: : : :| |::| |   :: |:|| | :|   :  :  | :
Qy    842 VGEEVISTLQILPTVREDSGFFSCHAINSYGEDRGIIQLTVQEPPDPPEIEI-KDVKART 900

Db    335 ITVTWDSG-NPD-PVSYYVIEYKSKSQD-GPYQIKEDITT--TRYSIGGLSPNSEYEIWV 389
          ||: |  | : : |:: | || |:||:   : |   |::      :|  : |:| | |::
Qy    901 ITLRWTMGFDGNSPITGYDIECKNKSDSWDSAQRTKDVSPQLNSATIIDIHPSSTYSIRM 960

Db    390 SAVNSIGQGPPSESVVTRTGEQAPASAPRNVQARMLSATTMIVQWEEPV-E-PNGLIRGY 447
          | | || : ||: :   : | || ::|::|:   :|: :: | |  |     ||:||||
Qy    961 YAKNRIGKSEPSNELTITADEAAPDGPPQEVHLEPISSQSI?VTWKAPKKHLQNGIIRGY 1020

Db    448 RV-Y--YTMEPEHPVGNWQKHNVDDSLLTTVGSLLEDETYTVRVLAFTSVGDGPLSDPIQ 504
          :: |  |:   :            || : |: :|     | : | |    | || |: |
Qy   1021 QIGYREYSTGGNIQFNIISVDTSGDSEVYTLDNLNKFTQYGLVVQACNRAGTGPSSQEII 1080

Db    505 VKTQQGVPG-QPMNLRAEAKSETSIGLSWSA-PRQ--ESVIK-YELLFREGDRGREVG-- 557
            | : ||:  | |::| | |  ||::|||: :::  ::::  : ::::      |:|
Qy   1081 TTTLEDVPSYPPENVQAIATSPESISISWSTLSKEALNGILQGFRVIYWANLMDGELGEI 1140

Db    558 RTFDPT-TAFVVEDLKPNTEYAFRLAARSPQGLGAFTAVVRQRTLQAISPKNFKVKMIM- 615
          : :  |  :: :: |   |:|:::: | :  | |  :  :  || : ::     ||
Qy   1141 KNITTTQPSLELDGLEKYTNYSIQVLAFTRAGDGVRSEQIFTRTKEDVPGPPAGVKAAAA 1200

Db    616 KTS-VLLSW 623
           :| |::||
Qy   1201 SASMVFVSW 1209


RESULT   13
ID   US-08-752-307B-13     STANDARD;     PRT;   596 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 13, Application US/08752307B
XX
CC    Sequence 13, Application US/08752307B
CC    Patent No. 5952171
CC     GENERAL INFORMATION:
CC       APPLICANT:  McCarthy, Sean A.
CC       APPLICANT:  Gearing, David P.
CC       APPLICANT:  Levinson, Douglas A.
CC       TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENES
CC       TITLE OF INVENTION:  ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Fish & Richardson, P.C.
CC         STREET:  225 Franklin Street
CC         CITY:  Boston
CC         STATE:  MA
CC         COUNTRY:  US
CC         ZIP:  02110-2804
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  Windows95
CC         SOFTWARE:  FastSEQ for Windows Version 2.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/752,307B
CC         FILING DATE:  19-NOV-1996
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Meiklejohn, Ph.D., Anita L.
CC         REGISTRATION NUMBER:  35,283
CC         REFERENCE/DOCKET NUMBER:  09404/020001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  617-542-5070
CC         TELEFAX:  617-542-8906
CC         TELEX:  200154
CC     INFORMATION FOR SEQ ID NO:  13:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  596 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   596 AA; 65965 MW; 1871085 CN;

  Query Match             4.0%;  Score 449;  DB 2;  Length 596;
  Best Local Similarity  23.8%;  Pred. No. 1.18e-20;
  Matches   134;  Conservative  138;  Mismatches 260;  Indels  32;  Gaps  29;

Db     44 KITLNCEARGNPSPHYRWQLNGSDIDTSLDHRYKLNGGNLIVINPNRNWDTGSYQCFATN 103
          :: | | | |:| | ||| |:  ::   |  |:       |:: |  |  |:||| |  :|
Qy    241 RVELPCKALGHPEPDYRW-LKD-NMPLELSGRFQKTVTGLLIEN-IRPSDSGSYVCEVSN 297

Db    104 SLGTIVSREAKLQFAYLENFKSRMRSR-VSVREGQGVVLLCGPPPHSGELSYAWVFNEYP 162
            ||      ::| :   : :|: : :| |     |  | | |:     : :   :| : :
Qy    298 RYGT-AKVIGRL-YV-KQPLKATISPRKVKSSVGSQVSLSCSVTG-TEDQELSW-YRNGE 352

Db    163 SFVEEDSRRFVSQETGHLYIAKVEPSDVGNYTCVVTSTVTNARVLGSPTPLVLRSDGVMG 222
          :    : |: : :  :| :  :  || | | | |      :|:        |   :  :::
Qy    353 ILNPGKNVRITGINHENLIMDHMVKSDGGAYQCFVRKDKLSAQDY-VQVVLEDGTPKIIS 411

Db    223 EYEPKIELQF-PETLP-AAKGSTVKLECFALGN-PVPQINWRRSDGMPFPTKIKLRKFNG 279
          :  |:     | :|    ||: :     :| : |:     :|   | : :  ::  :
Qy    412 AFSEKVVSPAEPVSLMCNVKGTPLPTITWTLDDDPILKGGSHRISQM-ITSEGNVVSY-- 468

Db    280 VLEIPNFQQEDTGSYECIAENSRGKNVARGRLTYYAKPYWVQLLKDVETAVEDSLYWECR 339
          |:|:: |  | | | | |:|| |  : ::|:    : |  :: :|:: :      |  ||
Qy    469 -LNISSSQVRDGGVYRCTANNSAGVVLYQARINVRG-PASIRPMKNITAIAGRDTYIHCR 526

Db    340 ASGKPKPSYRWLKNGDALVLEER-IQIEN-GALTIANL-NVSDSGMFQCIAENKHGLIYS 396
```

```
Qy    527 VIGYPYYSIKWYKNSNLLPFNHRQVAFENNGTLKLSDVQKEVDEGEYTCNVLVQPQLS-T 585

Db    397 SAELKVLASAPDFSRNPMKKMIQVQVGSLVILDCKPSASPRALSF-WKKGDTVVREQARI 455

Qy    586 SQSVHVTVKVPPFIQ-PFE-FPRFSIGQRVFIPCVVVSGDLPITITWQKDGRPIPGSLGV 643

Db    456 SLLN-D--GGLKIMNVTKADAGIYTCIAENQFGKANGTTQLVVTEPTRIILAPSNMDVAV 512

Qy    644 TIDNIDFTSSLRISNLSLMHNGNYTCIARNEAAAVEHQSQLIVRVPPKFVVQPRDQDGIY 703

Db    513 GESIILPCQVQHDPLLDIMFAWYFNGT-LTDFKKDGSHFEKVGGSSSGDLMIRNIQLKHS 571

Qy    704 GKAVI_NCSAEGYPVPTIVWK-FSKGAGVPQFQPIALN-GRIQVLSNGSLLIKHVVEEDS 761

Db    572 GKYVCMVQTGVDS-VSSAAELIVR 594

Qy    762 GYYLCKVSNDVGADVSKSMYLTVK 785


RESULT   14
ID   US-08-752-307B-11     STANDARD;     PRT;   612 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 11, Application US/08752307B
XX
CC    Sequence 11, Application US/08752307B
CC    Patent No. 5952171
CC     GENERAL INFORMATION:
CC       APPLICANT:  McCarthy, Sean A.
CC       APPLICANT:  Gearing, David P.
CC       APPLICANT:  Levinson, Douglas A.
CC       TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENES
CC       TITLE OF INVENTION:  ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Fish & Richardson, P.C.
CC         STREET:  225 Franklin Street
CC         CITY:  Boston
CC         STATE:  MA
CC         COUNTRY:  US
CC         ZIP:  02110-2804
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  Windows95
CC         SOFTWARE:  FastSEQ for Windows Version 2.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/752,307B
CC         FILING DATE:  19-NOV-1996
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Meiklejohn, Ph.D., Anita L.
CC         REGISTRATION NUMBER:  35,283
CC         REFERENCE/DOCKET NUMBER:  09404/020001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  617-542-5070
CC         TELEFAX:  617-542-8906
CC         TELEX:  200154
CC     INFORMATION FOR SEQ ID NO:  11:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  612 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   612 AA; 67666 MW; 1983167 CN;

  Query Match           4.0%;  Score 448;  DB 2;  Length 612;
  Best Local Similarity  24.9%;  Pred. No. 1.38e-20;
  Matches   149;  Conservative  138;  Mismatches 266;  Indels  45;  Gaps  41;

Db     35 EELSQPPTITQQSPKDYIVDPRENIVIQCEAKGKPPPSFSWTRNGTHFDIDKDAQVT-MK 93

Qy     18 EDLHSSLYFVNASLQEVVFASTTGTLVPCPAAGIPPVTLRWYLATGE-EIYDVPGIRHVH 76

Db     94 PNSGTL-VVNIM-NGVKAEAYEGVYQCTARNERGAAISNNIVIRPS-RSPLWTKEKLEPN 150

Qy     77 PN-GTLQIFPFPPSSFSTLIHDNTYYCTAENPSGKIRSQDVHIKAVLREP-YT-VRVEDQ 133

Db    151 HVREGDSLVLNCRPPVGLPP--PIIFWMDNAFQRLPQSERVSQGLNGDLYFSNVQPEDTR 208

Qy    134 KTMRGNVAVFKCIIPSSVEAYITVVSWEKDTV-SLVSGSRFLITSTGALYIKDVQNEDGL 192

Db    209 VDYICYARFNHTQTIQQKQPISVKVFSTKPVTERPPVLLTPMGSTSNKVELRGNVLLLEC 268

Qy    193 YNYRCITR--HRYTGETKQSNSARLFVSDP-ANSAPSILD--GFDHRKA-MAGQRVELPC 246

Db    269 IAAGLPTPVIRWIKEGGELPANRTFFENFKKTLKIIDVSEADSGNYKCTARNTLGSTHHV 328

Qy    247 KALGHPEPDYRWLKDNMPLELSGRF-QKTVTGLLIENIRPSDSGSYVCEVSNRYGTAKVI 305

Db    329 ISVTVKAAPYWITAPRNLVLSPGEDGTLICRANGNPKPSISWLTNGVPIAIAPEDPSRKV 388

Qy    306 GRLYVKQPLKATISPRKVKSSVGSQVSLSCSVTGTEDQELSWYRNGEILNPGKNVRITGI 365

Db    389 DGDTIIFSAVQERSSAVYQCNASNEYGYLLANAFVNV-LAE-PPRILTP-ANKLYQVIAD 445

Qy    366 NHENLIMDHMVKSDGGAYQCFVRKD-K-LSAQDYVQVVLEDGTPKIISAFSEKVVSP-AE 422

Db    446 SPALIDCAYFGSPKPEIEW-FRG--V-KGSILRGNEYVFHD-NG-T-LEIPVAQKDSTGT 498

Qy    423 PVSLM-CNVKGTPLPTITWTLDDDPILKGGSHRISQMITSEGNVVSYLNISSSQVRDGGV 481

Db    499 YTCVARNKLGKTQNEVQLEVKDPTMIIKQPQYKVIQRSAQASF-EC-VIKHDPTLIPTVI 556

Qy    482 YRCTANNSAGVVLYQARINVRGPASI-R-PMKNITAIAGRDTYIHCRVIGY-P-YY-SIK 536

Db    557 WLKDNNELPDDERFLV-GKDN-LTIMNV-TDKDDGTYTC-I-VNTTLDSVSASAVLTV 609

Qy    537 WYKNSNLLPFNHRQVAFENNGTLKLSDVQKEVDEGEYTCNVLVQPQL-STSQSVHVTV 593


RESULT   15
ID   US-08-752-307B-12     STANDARD;     PRT;   607 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 12, Application US/08752307B
XX
CC    Sequence 12, Application US/08752307B
CC    Patent No. 5952171
CC     GENERAL INFORMATION:
CC       APPLICANT:  McCarthy, Sean A.
CC       APPLICANT:  Gearing, David P.
CC       APPLICANT:  Levinson, Douglas A.
CC       TITLE OF INVENTION:  METHOD FOR IDENTIFYING GENES
CC       TITLE OF INVENTION:  ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Fish & Richardson, P.C.
CC         STREET:  225 Franklin Street
CC         CITY:  Boston
CC         STATE:  MA
CC         COUNTRY:  US
CC         ZIP:  02110-2804
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
```

```
CC         OPERATING SYSTEM:  Windows95
CC         SOFTWARE:  FastSEQ for Windows Version 2.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/752,307B
CC         FILING DATE:  19-NOV-1996
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Meiklejohn, Ph.D., Anita L.
CC         REGISTRATION NUMBER:  35,283
CC         REFERENCE/DOCKET NUMBER:  09404/020001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  617-542-5070
CC         TELEFAX:  617-542-8906
CC         TELEX:  200154
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  607 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   607 AA; 66494 MW; 1857918 CN;

  Query Match             3.5%;  Score 391;  DB 2;  Length 607;
  Best Local Similarity  25.8%;  Pred. No. 1.00e-16;
  Matches   129;  Conservative  121;  Mismatches 208;  Indels  42;  Gaps  39;

Db     58 LLACRARASPPATYRWKMNGTEMKLE-PGSRHQLVGGNL-VI-MNPTK-AQ--DAGVYQC 111
Qy     43 LVPCPAAGIPPVTLRWYLATGEEIYDVPGIRHVHPNGTLQIFPFPPSSFSTLIHDNTYYC 102

Db    112 LASNPVGTVVSREAILRFGFLQE-FS-K-EERDPVKAHEG-WGVMLPCNPPAHYPGLSYR 167
Qy    103 TAENPSGKIRSQDVHIK-AVLREPYTVRVEDQKTMRGNVAVFKCIIPSSVEA-YITVV-S 159

Db    168 WLLNEFPNFIPTDGRHFVSQTTGNLYI-ARTNASDLGNYSCLATSHMDFSTKSVFSKFAQ 226
Qy    160 WE-KDTVSLV-SGSR-FLITSTGALYIKDVQNEDGLYNYRCI-TRHR-YTGETRQSNSAR 214

Db    227 LNLAAEDTRLFAPSIKARFPAETYALVGQQVTLECFAFGNPVPRIKWRKVD-G-SLSPQW 284
Qy    215 L-FVSDPAN-SAPSILDGFDHRK-AMAGQRVELPCKALGHPEPDYRWLKDNMPLELSGRF 271

Db    285 TTAEPTLQIPSVSFEDEGTYECEAENSKGRDTVQGRIIVQAQPEWLKVIS-DTEADIGSN 343
Qy    272 QKTVTGLLIENIRPSDSGSYVCEVSNRYGTAKVIGRLYVK-QPLKATISPRKVKSSVGSQ 330

Db    344 LRWGCAAAGKPRPTVRWLRNGEPL-ASQN-RVEVLA-GDLRFSKLSLEDSGMYQC-VAEN 399
Qy    331 VSLSCSVTGTEDQELSWYRNGEILNPGKNVRITGINHENLIMDHMVKSDGGAYQCFVRKD 390

Db    400 KHGTIYASAELAVQALAPDFRLNPVRRLIPAARGGEILIPCQPRAAPKAVVLWSKGTE-I 458
Qy    391 KLSAQ-DYVQVVLEDGTPKIISAFSEKVVSPAEPVS-LM-CNVKGTPLPTITWTLDDDPI 447

Db    459 LVNSS-RVT--VTPDGTLII-RNISRSD--EGK-YTCFAENFMGKANSTGILSVRDATKI 511
Qy    448 LKGGSHRISQMITSEGNVVSYLNISSSQVRDGGVYRCTANNSAGVVLYQARINVRGPASI 507

Db    512 TLAPSSA.INLGDNLTLQCH 531
Qy    508 RPMKNITAIA-GRDTYIHCR 526


Search completed: Tue Apr 25 13:49:26 2000
Job time : 41 secs.
```